REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
06/03/2011

2. REPORT TYPE
Final Report

3. DATES COVERED *(From - To)*
01/29/2008 - 03/03/2011

4. TITLE AND SUBTITLE
Uncertainties on Networks

5a. CONTRACT NUMBER

5b. GRANT NUMBER
N00014-08-1-0464

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Cressie, Noel A.

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The Ohio State University
Department of Statistics
1958 Neil Avenue
Columbus, OH 43210-1247

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
The Ohio State University
Office of Sponsored Programs
1960 Kenny Road
Columbus, OH 43210

10. SPONSOR/MONITOR'S ACRONYM(S)
OSP

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
60015815

12. DISTRIBUTION/AVAILABILITY STATEMENT
No limitations or restrictions are pertinent to this report.

13. SUPPLEMENTARY NOTES

20110613180

14. ABSTRACT
Models in space and space-time are essential for representing the battlespace. For example, they are used in estimating a dynamically evolving danger function or in predicting a waypoint in the presence of uncertainties. In netcentric warfare, the uncertainties reside not only in the variables at the network's nodes, but also in the presence or absence of network nodes and the dependencies between the nodes. This occurs when a node may only be operational intermittently or when the enemy's network is unknown, apart from a few obvious nodes. In this research, we incorporate spatial and spatio-temporal dependencies into the analysis of network data.

15. SUBJECT TERMS
Acyclic directed graph (ADG); chain graph; Markov random fields; level sets; spatial statistics; spatio-temporal statistics.

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON / 19b. TELEPHONE NUMBER *(Include area code)*
U	U	U	UU	11	Kristen Mosier / 614-292-6847

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

DTIC® DEFENSE TECHNICAL INFORMATION CENTER

Information for the Defense Community

DTIC® has determined on 13 / JUN / 2011 that this Technical Document has the Distribution Statement checked below. The current distribution for this document can be found in the DTIC® Technical Report Database.

☒ **DISTRIBUTION STATEMENT A.** Approved for public release; distribution is unlimited.

☐ **© COPYRIGHTED**; U.S. Government or Federal Rights License. All other rights and uses except those permitted by copyright law are reserved by the copyright owner.

☐ **DISTRIBUTION STATEMENT B.** Distribution authorized to U.S. Government agencies only (fill in reason) (date of determination). Other requests for this document shall be referred to (insert controlling DoD office)

☐ **DISTRIBUTION STATEMENT C.** Distribution authorized to U.S. Government Agencies and their contractors (fill in reason) (date of determination). Other requests for this document shall be referred to (insert controlling DoD office)

☐ **DISTRIBUTION STATEMENT D.** Distribution authorized to the Department of Defense and U.S. DoD contractors only (fill in reason) (date of determination). Other requests shall be referred to (insert controlling DoD office).

☐ **DISTRIBUTION STATEMENT E.** Distribution authorized to DoD Components only (fill in reason) (date of determination). Other requests shall be referred to (insert controlling DoD office).

☐ **DISTRIBUTION STATEMENT F.** Further dissemination only as directed by (inserting controlling DoD office) (date of determination) or higher DoD authority.

Distribution Statement F is also used when a document does not contain a distribution statement and no distribution statement can be determined.

☐ **DISTRIBUTION STATEMENT X.** Distribution authorized to U.S. Government Agencies and private individuals or enterprises eligible to obtain export-controlled technical data in accordance with DoDD 5230.25; (date of determination). DoD Controlling Office is (insert controlling DoD office).

FINAL TECHNICAL REPORT
"UNCERTAINTIES ON NETWORKS"

Grant No.:	N00014-08-1-0464
Period of Award:	01/29/2008-03/03/2011
PI Name:	Noel Cressie
Address:	Department of Statistics The Ohio State University 1958 Neil Avenue Columbus OH 43210-1247
Phone Number:	614-688-0072
Fax Number:	614-292-2096
e-mail Address:	ncressie@stat.osu.edu

UNCERTAINTIES ON NETWORKS (PI: N. Cressie)

Objectives

Networks as models can be found in many disciplines, including biology, computer science, engineering, geography, mathematics, physics, sociology, and statistics. However, there are uncertainties associated with imperfect knowledge of a network's nodes and dependencies, as well as with noise-corrupted variables defined on the network. Networks have become important components of complex representations of reality and, when built into a hierarchical statistical modeling structure, they allow partitioning of joint probability distributions that seem unmanageable at first glance. Thus, a statistical approach to network analysis is natural from both a probabilistic and an inferential point of view. In this research, we study spatial and spatio-temporal networks through graph theory (e.g., Lauritzen, 1996; Cressie and Davidson, 1998). A chain graph is defined to be a combination of undirected graphs and acyclic directed graphs (ADGs), with the overall structure being guided by an ADG. The undirected parts account for the spatial dependence, the directed parts can be used to account for the temporal dependence, and the guiding ADG captures the spatio-temporal interactions.

Impact

Models in space and space-time are essential for representing the battlespace. For example, they are used in estimating a dynamically evolving danger function or in predicting a waypoint in the presence of uncertainties. In netcentric warfare, the uncertainties reside not only in the variables at the network's nodes, but also in the presence or absence of network nodes and the dependencies between the nodes. This occurs when a node may only be operational intermittently or when the enemy's network is unknown, apart from a few obvious nodes. In this research, we incorporate spatial and spatio-temporal dependencies into the analysis of network data.

References

Cressie, N. and Davidson, J.L. (1998). Image analysis with partially ordered Markov models. *Computational Statistics and Data Analysis*, **29**, 1-26.

Lauritzen, S.L. (1996). *Graphical Models*, Oxford University Press: Oxford, UK.

UNCERTAINTIES ON NETWORKS (PI: N. Cressie)

Personnel

Principal Investigator: Noel Cressie, PhD

Research Assistants: Lili Zhuang
Emily Kang
Rajib Paul
Aritra Sengupta
Jon Bradley

Award Participants: PI: male; GRAs: 2 females and 3 males

Technical Summary

Recently, a great amount of attention has been paid to random networks, which are widely used to represent complex relationships in many areas (e.g., World Wide Web communications, social studies, epidemic dynamics, molecular-evolution processes, etc.). According to Lauritzen (1996), a random network can be modeled through a mathematical graph defined as $\mathbf{G} = (\mathbf{V}, \mathbf{E})$, which consists of a finite set of nodes (or vertices), $\mathbf{V}$, and a set of edges, $\mathbf{E}$, where nodes represent individuals or objects, while edges specify their relationships.

Graphs can be further divided into different classes, according to the nature of their edges as well as the paths formed by edges. Our research focuses on one type of graph called a *chain graph*, made up of *undirected graphs* and *acyclic directed graphs* (ADGs or sometimes abbreviated as DAGs). ADGs consist of only directed edges without any cycles, and thus they can specify direct relations (e.g., conditional dependencies, causal relations) between variables defined on the graph's nodes; see Lauritzen (1996), Kolaczyk (2009), and Koski and Noble (2009) for further details.

In recent research on statistical-dependence modeling, Bayesian networks are widely used to characterize joint multivariate probability distributions, which can define properties of conditional independence or causal relations between variables in a complex process. In the research conducted under this grant, we incorporate spatial and spatio-temporal dependencies into the analysis of network data.

An explosion of ideas has been generated on dependence modeling based on networks (e.g., Friedman et al., 2000; Ellis and Wong, 2008). According to Koski and Noble (2009), a Bayesian network, BN $= (\mathbf{G}, p)$, can be modeled through an ADG, $\mathbf{G}$, and its probability distribution, p. The Erdös-Rényi model (E-R model) has been widely used in the past to capture the probability distributions of ADGs (Erdös and Rényi, 1959). This model belongs to the family of *exponential random graph models* (ERGM) (e.g., Hunter and Handcock, 2005), and it assumes equal and independent probabilities of having an edge between any pair of nodes within a graph (referred as "equal and independent assumptions"). The E-R model is also frequently used as a prior distribution for ADGs with discrete data. The main appeal of the E-R model is that it can lead to a closed-form posterior distribution (e.g., Ellis and Wong, 2008). However, in reality, the equal and independent assumptions of the E-R model are not realistic, especially for high-dimensional networks. Furthermore, its sufficient statistic captures only one property of a random graph, namely the number of edges; all the other important properties, such as the directions of edges, the patterns formed by the edges are ignored.

In what follows, we consider more general ADGs, based on the level-set definition proposed by

Cressie and Davidson (1998). We develop a sequential-modeling strategy, through which we can capture the probability distributions of ADGs, but we avoid strong assumptions such as the equivalent and independent assumptions. Furthermore, our level-set model allows more graphical information to be used; for example, we consider not only the number of edges, but also certain structure of the ADG, including levels of the ADG, connections between levels (definition of "levels" and "connections" will be given later), directions of edges between nodes, etc. Based on our level-set modeling strategy, we also develop an algorithm to generate ADGs efficiently.

We introduce the following notation

- $\mathbf{G}$ denotes an ADG, and $\mathbf{G} = (\mathbf{V}, \mathbf{E})$.
- $\mathbf{V}$ denotes the set of finite nodes in $\mathbf{G}$; that is, $\mathbf{V} = \{v_1, ..., v_n\}$, where n is the total number of nodes and n is given.
- $\mathbf{E}$ denotes the set of directed edges in $\mathbf{G}$:

$$\mathbf{E} = \{(v_i, v_j) : \text{ there is a directed edge from } v_i \text{ to } v_j; v_i, v_j \in \mathbf{V}\}. \tag{1}$$

- $\mathrm{ch}(v_i)$ denotes the children of node v_i; that is, for $v_i \in \mathbf{V}$,

$$\mathrm{ch}(v_i) = \{v_j \in \mathbf{V} : (v_i, v_j) \in \mathbf{E}\}. \tag{2}$$

- $\mathrm{pa}(v_i)$ denotes the parents of node v_i; that is, for $v_i \in \mathbf{V}$,

$$\mathrm{pa}(v_i) = \{v_j \in \mathbf{V} : (v_j, v_i) \in \mathbf{E}\}. \tag{3}$$

- $\mathbf{V}_{min}$ denotes the set of vertices with no parents; that is,

$$\mathbf{V}_{min} = \{v_i \in \mathbf{V} : \mathrm{pa}(v_i) = \varnothing\}. \tag{4}$$

- $\mathrm{covr}(\mathbf{B})$ denotes the cover of a subset of nodes $\mathbf{B} \subset \mathbf{V}$, which is the subset of nodes that are not in $\mathbf{B}$ but whose parents are all in $\mathbf{B}$ (Cressie and Davidson, 1998); that is,

$$\mathrm{covr}(\mathbf{B}) = \{v_i \in \mathbf{V} : \mathrm{pa}(v_i) \subset \mathbf{B} \text{ and } v_i \notin \mathbf{B}\}. \tag{5}$$

Notice that the definition of the cover of a subset of nodes is different from the Markov blanket (e.g., Pearl, 1988); for a set of nodes, the Markov blanket consists of their children, their parents, as well as their children's other parents. In other words, the Markov blanket contains all the variables that shield the subset of nodes from the rest of the network. However, $\mathrm{covr}(\mathbf{B})$ only includes certain descendants: $\mathrm{covr}(\mathbf{B}) \subset \mathrm{ch}(\mathbf{B})$.

From Cressie and Davidson (1998), an ADG with a finite number of nodes has level sets $\mathbf{L} = \{\mathbf{L}_0, ..., \mathbf{L}_d\}$, formed by a specific partition of the ADG that can be specified recursively as,

$$\mathbf{L}_i \equiv \begin{cases} \mathbf{V}_{min}, & \text{if } i = 0; \\ \mathrm{covr}(\cup\{\mathbf{L}_k : k = 0, ..., i-1\}), & \text{if } 0 < i \leq d, \end{cases} \tag{6}$$

where $(d+1)$ is the total number of level sets. For an ADG with n nodes, it is straightforward to see that $1 \leq d+1 \leq n$. The important properties of level sets can be summarized below (Cressie and Davidson, 1998):

1. Every node of an ADG should belong to one and only one level set; specifically,

$$\mathbf{L}_i \cap \mathbf{L}_j = \varnothing \text{ for } i \neq j = 0, ..., d. \tag{7}$$

In other words, the $(d+1)$ level sets, $\mathbf{L} = \{\mathbf{L}_0, ..., \mathbf{L}_d\}$, together form a $(d+1)$-nonempty-partition of the ADG.

2. Every node in a non-minimal level set should have at least one parent from its adjacent level set that is of lower order; that is, for any $v \in \mathbf{L}_i$, $0 < i \leq d$, then there exists a node $u \in \mathbf{L}_{i-1}$, such that $u \in \text{pa}(v)$.

3. The directed edges can only go from nodes in lower-order level sets to nodes in higher-order level sets; that is, if $v, u \in \mathbf{V}$, $v \in \mathbf{L}_i$, $0 < i \leq d$, and $u \in pa(v)$, then $u \in \cup\{\mathbf{L}_k : k = 0, ..., i-1\}$.

4. The nodes in the same level set should be independent; that is, there are no directed edges within any level set. In other words, if $v, u \in \mathbf{V}$ and $v, u \in \mathbf{L}_i$, $i = 0, ..., d$, and $v \neq u$, then there should be no directed edge between v and u.

5. If $v \in \mathbf{L}_i$, $0 < i \leq d$, then there should be a path of length i from a given node $u \in \mathbf{L}_0$ to v. Furthermore, no path to v can be longer than length i.

6. The maximum length of a path in an ADG with $(d+1)$ level sets, $\mathbf{L} = \{\mathbf{L}_0, ..., \mathbf{L}_d\}$, is d.

According to the definition and properties mentioned above, we notice that different ADGs can give rise to the same level-sets structure; however, given an ADG, the level-sets structure should be unique. This is an important property that differentiates the notion of level sets from other modern graph-partition strategies (e.g., partitions to obtain minimal edges among partitions but maximal edges within partitions; see Newman, 2004). Those types of graph partitions are not unique for a given ADG.

Figure 1 shows an ADG with level-sets structure satisfying all the properties mentioned above. For example, nodes v_1 and v_2 are in the minimal level set $\mathbf{L}_0$, because they have no parents; there is no directed edge within each level set; directed edges always go from lower-order level sets to higher-order level sets, and so forth.

From the definitions and properties of level sets given above, we can see that the level-sets structure of an ADG involves much more graphical information than just the number of edges found in the E-R model. In order to specify the structure between level-sets, we introduce the connection matrix, but we first need to define the adjacency matrix.

From Lauritzen (1996), we can use an *adjacency matrix*, $\mathbf{Y} \equiv [y_{ij}]_{n \times n}$, to uniquely specify the structure of an ADG with n nodes:

$$y_{ij} \equiv \begin{cases} 1, & \text{if there is a directed edge from } v_i \text{ to } v_j, \text{ where } v_i, v_j \in \mathbf{V}; \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

Similarly, we can define a *connection matrix*, $\mathbf{C} \equiv [c_{kl}]_{(d+1)\times(d+1)}$, to specify the structure between level sets. Consider a given ADG with $(d+1)$ level sets, $\mathbf{L} = \{\mathbf{L}_0, ..., \mathbf{L}_d\}$. If there is at least one directed edge going from one of the nodes in level set $\mathbf{L}_i$ to one of the nodes in level set $\mathbf{L}_j$,

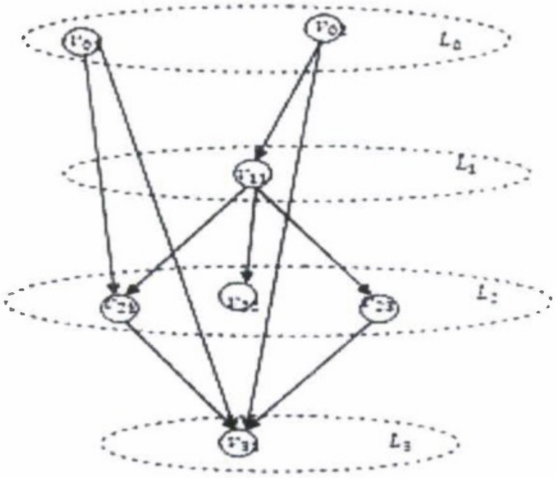

Figure 1: An ADG with 7 nodes and 4 level sets

$i < j$, then we say that there is a directed connection going from $\mathbf{L}_i$ to $\mathbf{L}_j$. Otherwise, if there is no directed edge between nodes in two different level sets within an ADG, we say that there is no connection between the two level sets. Thus, we define the connection matrix of level sets, $\mathbf{L} = \{\mathbf{L}_0, ..., \mathbf{L}_d\}$, as a $(d+1) \times (d+1)$ matrix, $\mathbf{C} \equiv [c_{kl}]_{(d+1)\times(d+1)}$, as follows:

$$c_{k+1,l+1} \equiv \begin{cases} 1, & \text{if there is a directed connection from } L_k \text{ to } L_l, \text{ where } L_k, L_l \in \mathbf{L}; \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

For example, the connection matrix $\mathbf{C}$ of the ADG with seven nodes and four level sets in Figure 1 can be written as,

$$\mathbf{C} = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{10}$$

Now we shall discuss modeling strategies for ADGs. As mentioned before, the ERGM family is popular for modeling ADGs. A typical ERGM defines the probability of an ADG as (Hunter and Handcock, 2005):

$$\mathrm{P}(\mathbf{G}|\boldsymbol{\Theta}) = \frac{\exp[\boldsymbol{\Theta}^T \mathrm{g}(\mathbf{G})]}{c(\boldsymbol{\Theta})}, \tag{11}$$

where $\boldsymbol{\Theta}$ is a vector of parameters; $\mathrm{g}(\mathbf{G})$ is a vector of graph statistics that is sufficent for (11); and $c(\boldsymbol{\Theta})$ is the normalizing constant. For example, the E-R model is a specific case of an ERGM defined as:

$$\mathrm{P}(\mathbf{G}|\theta) \propto \exp[-\theta|\mathbf{G}|], \tag{12}$$

where $|\mathbf{G}|$ is the number of edges in $\mathbf{G}$, $\theta > 0$, and e^{θ} is interpreted as the probability of having an edge between any pair of nodes in the ADG $\mathbf{G}$. Therefore, the E-R model implies that the probability of having an edge between each pair of nodes is equal and independent within a graph. Although the ERGM family allows inclusion of other graphical structures, research on what type of graphical statistics can be used in the ERGM is still in its infancy. Furthermore, the ERGM family is difficult to apply to high-dimensional networks.

We shall propose level-sets models below to avoid these limitations of ERGMs. Rather than directly modeling the joint probability distribution of every individual node in the graph, our strategy is to first model the probability distribution of the unique level-sets structure of an ADG; then, we model the probability distributions of ADGs *conditional* on its level-sets structure. This uses information on the children's and parents' directed edges contained in the level-set structure. Also, this conditional-probability modeling strategy helps us avoid the strong assumptions made in defining ERGMs.

We define the level-sets model as follows:

$$P(\mathbf{G}|\mathbf{\Theta}) = P(\mathbf{G}|\mathbf{C},\mathbf{\Theta})P(\mathbf{C}|\mathbf{L},\mathbf{\Theta})P(\mathbf{L}|\mathbf{V},\mathbf{\Theta})P(\mathbf{V}|\mathbf{\Theta}) \tag{13}$$

Since the adjacency matrix $\mathbf{Y}$ and the ADG $\mathbf{G}$ are in one-to-one correspondence, we also can write equation (13) as

$$P(\mathbf{Y}|\mathbf{\Theta}) = P(\mathbf{Y}|\mathbf{C},\mathbf{\Theta})P(\mathbf{C}|\mathbf{L},\mathbf{\Theta})P(\mathbf{L}|\mathbf{V},\mathbf{\Theta})P(\mathbf{V}|\mathbf{\Theta}) \tag{14}$$

where, $\mathbf{\Theta}$ is a vector of parameters (e.g., Ellis and Wong, 2008).

Based on the level-sets model (13), we develop an associated algorithm that can efficiently generate ADGs. Compared to the E-R model, our level-sets model is appealing as a flexible prior distribution for Bayesian inference on ADGs. Zhuang and Cressie (2011) show how this algorithm can be used in Bayesian inference for multivariate distributions defined by ADGs and eventually chain graphs.

References

Cressie, N. and Davidson, J. (1998). Image analysis with partially ordered Markov models. *Computational Statistics and Data Analysis*, 29:1–26.

Ellis, B. and Wong, W. (2008). Learning causal Bayesian network structures from experimental data. *Journal of the American Statistical Association*, 103:778–789.

Erdös, P. and Rényi, A. (1959). On random graphs, I. *Publicationes Mathematicae Debrecen*, 6:290–297.

Friedman, N., Linial, M., Nachman, I., and Pe'er, D. (2000). Using Bayesian networks to analyze expression data. *Journal of Computational Biology*, 7:601–620.

Hunter, D. and Handcock, M. (2005). Inference in curved exponential family models for networks. *Journal of Computational and Graphical Statistics*, 15:565–583.

Kolaczyk, E. (2009). *Statistical Analysis of Network Data: Methods and Models.* Springer, New York.

Koski, T. and Noble, J. (2009). *Bayesian Networks: An Introduction.* Wiley, Chichester.

Lauritzen, S. (1996). *Graphical Models.* Oxford University Press, Oxford, UK.

Newman, M. (2004). Detecting community structure. *European Physical Journal B*, 38:321–330.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems.* Morgan Kaufmann, San Mateo, CA.

Zhuang, L. and Cressie, N. (2011). A class of priors on Bayesian ADG networks. In preparation.

UNCERTAINTIES ON NETWORKS (PI: N. Cressie)

Results

Listed below are the publications and presentations related to the grant.

Publications: Refereed

Cressie, N. and Kapat, P. (2008). Some diagnostics for Markov random fields. *Journal of Computational and Graphical Statistics*, **17**, 726-749.

Huang, C., Yao, Y., Cressie, N., and Hsing, T. (2009). Multivariate intrinsic random functions for cokriging. *Mathematical Geosciences*, **41**, 887-904.

Cressie, N., Shi, T., and Kang, E.L. (2010). Fixed rank filtering for spatio-temporal data. *Journal of Computational and Graphical Statistics*, **19**, 724-745.

Huang, C., Hsing, T., Cressie, N., Ganguly, A.R., Protopopescu, V.A., and Rao, N.S. (2010). Bayesian source detection and parameter estimation of a plume model based on sensor network measurements (with discussion). *Applied Stochastic Models in Business and Industry*, **26**, 331-361.

Kang, E.L., Cressie, N., and Shi, T. (2010). Using temporal variability to improve spatial mapping with application to satellite data. *Canadian Journal of Statistics*, **38**, 271-289.

Paul, R. and Cressie, N. (2011). Lognormal block kriging for contaminated soil. *European Journal of Soil Science*, **62**, 337-345.

Sain, S., Furrer, R., and Cressie, N. (2011). Combining ensembles of regional climate model output via a multivariate Markov random fields model. *Annals of Applied Statistics*, **5**, 150-175.

Braverman, A., Cressie, N., and Teixeira, J. (2011). A likelihood-based comparison of temporal models for physical processes. *Statistical Analysis and Data Mining*, forthcoming.

Huang, C., Hsing, T., and Cressie, N. (2011). Spectral density estimation through a regularized inverse function. *Statistica Sinica*, forthcoming.

Kang, E.L., and Cressie, N. (2011). Bayesian inference for the spatial random effects model. *Journal of the American Statistical Association*, forthcoming.

Zhuang, L. and Cressie, N. (2011). Spatio-temporal modeling of Sudden Infant Death Syndrome data. *Statistical Methodology*, forthcoming.

Publications: Nonrefereed

Sain, S.R., Furrer, R., and Cressie, N. (2009). Combining regional climate model output via a multivariate Markov random field model, in *Proceedings of the 56th Session of the International Statistical Institute*, Lisbon, Portugal, 1375-1382.

Cressie, N. and Kang. E. L. (2010). High-resolution digital soil mapping: Kriging for very large datasets, in *Proximal Soil Sensing*, eds. R. Viscarra-Rossel, A.B. McBratney, and B. Minasny. Springer, Dordrecht, 49-63.

Cressie, N. and O'Donnell, D. (2010). Statistical dependence in stream networks. A comment on "A moving average approach for spatial statistical models for stream networks" by J. Ver Hoef and E. Peterson. *Journal of the American Statistical Association*, **105**, 18-21.

Publications: Submitted

Cressie, N., Assunção, R., Holan, S.H., Levine, M., Nicolis, O., Zhang, J., and Zou, J. Dynamical random-set modeling of concentrated precipitation in North America. *Statistics and its Interface*, under journal review.

Huang, C., Hsing, T., and Cressie, N. Nonparametric estimation of the variogram and its spectrum, under revision for *Biometrika.*

Kang, E.L., Cressie, N., and Sain, S. Combining outputs from the NARCCAP regional climate models using a Bayesian hierarchical model, under revision for *Applied Statistics.*

Li, H., Calder, C.A., and Cressie, N. One-step estimation of spatial dependence parameters: Properties and extensions of the APLE statistic, under revision for *Journal of Multivariate Analysis.*

Nguyen, H., Cressie, N., and Braverman, A. Spatial statistical data fusion for remote-sensing applications, under journal review.

Presentations

April 2008, presented an invited paper at Workshop on Research Directions in Information Integration, Monterey, CA; "Uncertainties on networks."

April 2008, invited seminar speaker, Department of Statistics, University of California at Los Angeles; "Spatial prediction for massive datasets."

June 2008, invited speaker in the Science Visitors and Colloquium Program, Jet Propulsion Laboratory, Pasadena, CA; "Thinking statistically about analyzing global environmental datasets."

April 2009, presented the Shumway Lecture, Statistical Science Symposium, University of California, Davis; "Fixed rank filtering for spatio-temporal data."

June 2009, invited seminar speaker, CSIRO Division of Mathematical and Information Science (CMIS), Perth, Australia; "Using temporal variability to improve spatial mapping of satellite data."

July 2009, invited seminar speaker, CMIS, CSIRO, Brisbane, Australia; "Using temporal variability to improve spatial mapping of satellite data."

July 2009, presented a keynote address to 2009 International Symposium in Statistics, St. John's, Newfoundland, Canada; "Using temporal variability to improve spatial mapping."

August 2009, presented the 2009 Fisher Lecture at Joint Statistical Meetings (JSM), Washington, DC; "Where, when, and then why."

August 2009, presented an Introductory Overview Lecture at the Joint Statistical Meetings, Washington, DC; "Spatial statistical thinking."

August 2009, co-authored a contributed paper (with L. Zhuang) at JSM, Washington, DC; "Posterior distribution on networks."

August 2009, invited seminar speaker at IBM Thomas J. Watson Research Center, Yorktown Heights, NY; "Using temporal variability to improve spatial mapping of satellite data."

December 2009, presented 5 hours of lectures at Statistical Society of Australia Space-Time Modelling Symposium, Canberra, Australia; "Spatio-temporal statistics."

January 2010, presented 5 hours of lectures at Suisse Romand Troisième Cycle, Les Diablerets, Switzerland; "Spatio-temporal statistics."

February 2010, invited seminar speaker, Department of Statistics and Operations Research, University of North Carolina, Chapel Hill; "Spatio-temporal statistics: Fixed rank filtering applied to satellite data."

February 2010, presented an invited paper (with S. Holan) at SAMSI Workshop on Climate Change, Research Triangle Park, NC; "Computation, visualization, and dimension reduction in spatio-temporal modeling."

June 2010, presented the Opening Keynote Address at METMA V, International Workshop on Spatio-Temporal Modeling, Santiago de Compostella, Spain; "Dynamical temporal modeling of spatial fields."

August 2010, presented an invited paper at JSM, Vancouver, Canada; "The spatial random effects model and its role in spatial and spatio-temporal statistics."

August 2010, co-authored a contributed paper (with L. Zhuang) at JSM, Vancouver, Canada; "Spatio-temporal modeling of Sudden Infant Death Syndrome data."

August 2010, attended SAMSI opening workshop on Complex Networks, Research Triangle Park, NC.

October 2010, presented the Opening Keynote Address at SAMSI Spatial Transition Workshop, Research Triangle Park, NC; "Spatio-temporal statistical modeling."

December 2010, presented an invited paper at Workshop on Statistical Frontiers, 2010 ISS, Taipei, Taiwan; "Spatio-temporal statistical modeling."

April 2011, presented an invited paper (remotely) at National Center for Ecological Analysis and Synthesis (NCEAS) Workshop on Spatial Statistical Models for Stream Networks, Santa Barbara, CA; "Spatio-temporal statistical modeling."

April 2011, presented an invited paper at NIMBioS Workshop on Synchrony in Biological Systems, Knoxville, TN; "Spatio-temporal statistics."

Honors, Awards, and Prizes

Fellow, Spatial Econometrics Association, 2008

Science Visitor, Jet Propulsion Laboratory (NASA), 2008

Commonwealth Scientific and Industrial Research Organization, Distinguished Visiting Scientist, 2009

RA Fisher Award and Lectureship, Committee of Presidents of Statistical Societies (COPSS), 2009; www.amstat.org/meetings/jsm/2009/webcasts/wmv/2009fisher_Lecture.wmv

IBM Faculty Award, 2009

Member of Thomson Reuters 2010 ISI Highly Cited Researchers in Mathematics (out of a total of 356 in the world)

From ISI Web of Knowledge, March 2011: h-index = 32.